

Development and Validation of the ‘iCAN!’ - A Self-administered Questionnaire Measuring Outcomes/Competences and Professionalism of Medical Graduates

Ioannis DK Dimoliatis^{1,*}, Georgios N Lyrakos², Xanthippi Tseretopoulou¹, Theodoros Tzamalīs¹, George Bazoukis¹, Alexis Benos³, Charalambos Gogos⁴, Konstantinos Malizos⁵, Ioannis Pneumatikos⁶, Kyriaki Thermos⁷, Eleni Kaldoudi⁶, Margaret Tzaphlidou¹, Iordanis N Papadopoulos², Eleni Jelastopulu⁴

¹University of Ioannina Medical School, Ioannina, Greece

²National and Capodestrian University of Athens Medical School, Athens, Greece

³Aristotle University of Thessaloniki Medical School, Thessaloniki, Greece

⁴University of Patras Medical School, Patras, Greece

⁵University of Thessaly Medical School, Larissa, Greece

⁶Democritus University of Thrace, Medical School, Alexandroupolis, Greece

⁷University of Crete Medical School, Heracleion, Greece

*Corresponding Author: idimolia@uoi.gr

Copyright © 2014 Horizon Research Publishing All rights reserved.

Abstract The Tuning-Medicine Project produced a set of ‘level one’ and ‘level two’ learning outcomes/competences to be met by European medical graduates. In the learner-centered era self-assessment becomes more and more important. Our aim was to develop a self-completion questionnaire (‘iCAN!’) evaluating graduates’ learning outcomes. The Tuning ‘level two’ learning outcomes were transformed to a 104 closed items questionnaire, tested with a convenient sample of 512 graduates from the seven Greek medical schools during the 2009–2011 graduation cohorts, 21 practitioners, and seven different undergraduate student groups. Cronbach alpha, factor analysis, and mean score comparisons were used to check internal consistency, construct validity, and sensitivity respectively. Question mean scores were used to differentiate weak and strength areas of competence. Among graduates, all overall alphas were >0.95 and all subscale alphas higher than expected, indicating an at least acceptable internal consistency. Factor analysis produced one factor per ‘level one outcome’, except for ‘practical procedures’ which was split in two meaningful factors. Overall mean score was 44.4%, 52.2%, 61.2%, 73.4%, 81.4% among 2nd-year, 4th-year, 6th-year students, graduates, and practitioners respectively ($p < 0.001$); improvement through progressively matured groups suggest good sensitivity, indicating also good responsiveness. Among graduates, question mean scores ranged from very weak (blood transfusion, 39%) to very strength (measuring blood pressure, 95%), indicating good differentiation of strengths and weaknesses. A consistent, well-structured, and sensitive version of a questionnaire is hereby released for graduates’ abilities and professionalism self-assessment and differentiation of strengths and weaknesses that could be

used for informed SWOT policy.

Keywords Competence, Graduate Medical Education, ‘iCAN!’, Outcome Based Education, Questionnaire Development, Self Assessment, Validation

1. Introduction

Medical Education moved successively from delivering the content by medical teachers (content-based) to objectives-based, competence-based, and finally outcome-based education with special guides and journal themed issues [1-13]. Many schools worldwide have accepted the outcome-based education philosophy [14-22], and global minimum essential requirements have been defined [23-24]. It has penetrated even legislation [25]. In Europe, landmarks of this trend could be considered the European Core Curriculum: the Students’ Perspective [26-27] and mainly ‘The Tuning Project’ (medicine): Learning Outcomes/Competences for Undergraduate Medical Education in Europe, i.e. the teachers’ perspective, developed by the thematic network of ninety-three European medical schools on Medical Education in Europe (MEDINE), and accepted by the European Union [2]. These documents indicate a rather broad agreement on what medical graduates should be able to do at their graduation.

In any school’s curriculum assessment, one should listen to the students [28] who are perhaps the hardest judges [29]. Though relevant attempts have been published [24, 30, 31] no tool exists for measuring the end-product of a medical

curriculum. The need for measuring the educational environment during the period of undergraduate or postgraduate medical studies has led to a series of well-validated measures, where trainees' self-perceptions are used for measuring the educational environment [32-37]. Based on our experience of translating them into the Greek language [38,39] we constructed the '*iCAN!*' instrument, a self-completion questionnaire revealing graduates' perceptions on what they are able to do upon graduation [40]. This study's objective is to present the development and validation process.

2. Materials and Methods

2.1. Learning Outcomes for Undergraduate Medical Curriculum: The Content of the '*iCAN!*' Tool

Learning outcomes are usually specified with a hierarchy of levels, with a top level consisting of large domains of learning, within each of which subsidiary outcomes are described with increasing levels of granularity [7]. The European Core Curriculum was structured in nine domains ('level one' learning outcomes), each of which was comprised of four to twenty 'level two' learning outcomes, 92 in total [26]. For example, the first 'level two' learning outcome of the first 'level one' learning outcome states "Graduates should be able to take a detailed and relevant history". The Tuning Learning Outcomes were structured in domains (sixteen 'level one' learning outcomes, twelve for medical abilities and four for professionalism), each of which was comprised of up to thirteen 'level two' learning outcomes, 95 in total (69 for medical competences and 26 for professionalism). For example, the first 'level two' learning outcome of the first 'level one' learning outcome states "Graduates in medicine will have the ability to take a history" (Table 1).

Based mainly on the 'Tuning Learning Outcomes' [2] and supplemented by some important ones from the European Core Curriculum [26] the '*iCAN!*' questionnaire was developed, by transforming the Tuning 'level two' learning outcomes in a self-assessment tool. This instrument could be considered as a step further of the work done by the MEDINE group. The development of the tool followed three phases, each ending in a version of the questionnaire: the piloted initial version by undergraduate students, the tested version (v0) with graduates, and the hereby released final version (v1). Each version was built upon the experience of the previous one.

2.2. Development of the '*iCAN!*' Tool: The Initial Version

Firstly, the 'Tuning Learning Outcomes' were translated into the Greek language [41]. Then, according to the rules "each Tuning 'level two' learning outcome should be transformed into a single question" and "the whole tool should be kept as short as possible", an initial version of the questionnaire was created, taking into consideration "a

medical graduate completing it immediately after graduation" as a third rule.

There were 95 Tuning 'level two' learning outcomes; thus, the questionnaire should consist of 95 questions. However, since several Tuning 'level two' learning outcomes were bipartite, e.g. "subcutaneous and intramuscular injection" (one could be competent to the one but not to the other), the total items of the questionnaire increased. All bipartite 'level two' learning outcomes were split in two separate questions. In addition, the eleventh Tuning 'level one' learning outcome "Ability to apply scientific principles, method and knowledge to medical practice and research", not expanded to 'level two' learning outcomes, was also split in two separate questions, one relating to practice and one to research. On the other hand, the two 'level two' learning outcomes 16.1 "appreciation of diversity and multiculturalism" and 16.2 "understanding of cultures and customs of other countries" were considered too similar and thus combined in one question, while the 'level two' learning outcome 3.5 "advanced life support" was considered not suitable for the undergraduate level and thus omitted. Finally, three European Core Curriculum outcomes (diagnostic and therapeutic options available within other healthcare professions; STEEEP principle; non-verbal communication), not matching any Tuning learning outcome, were considered very important, therefore three more questions were added. In conclusion, a 122 items questionnaire was created (the initial version), accompanied by an open question ("Comments"). Since graduates are asked whether they are competent in all these outcomes, all questions started with the verb "I can", hence the name of the questionnaire; No negative question was formulated.

The initial version of the questionnaire was piloted in 15 third-year, 14 fifth-year and 18 sixth-year medical students at the Ioannina University, Greece. They were also asked whether they perceived the questions to be conceivable and/or duplicated. In regard to their input, the questions were restricted to 104 and subsequently randomly re-ordered using Microsoft Excel function 'rand()'. Though answering a long questionnaire is not an easy task, no other questions were eliminated in order to include all Tuning 'level two' learning outcomes. A last question (q105) that was raised during the first phase was added to investigate whether graduates attributed their abilities to their school's curriculum; it is worth to note that this is not an outcome question but an educational process question and its interpretation should be similar to educational environment measures such as DREEM, PHEEM etc. Finally, following Whittle et al [42], the open question of the initial version was replaced by "If you could change one thing in your school, what would this be?" [43].

Respondents were asked to choose one of six response options in a Likert scale (strongly agree, agree, rather agree, rather disagree, disagree, strongly disagree), without allowing uncertainty in the middle that causes loss of information [38]. In addition, offering participants six instead of five options, it is closer to the ideal 'five to nine' options, congruent with both respondent preferences and

reliability statistics, and at least somehow averts the ‘central tendency bias’ that results in loss of reliability and sensitivity. [44]

Thus, the version ‘iCAN!v0’ was created to be tested with a group of graduates, measuring their perceptions upon their medical abilities and professionalism (Table 1).

Table 1. The ‘Tuning learning outcomes’ and the related questions of the ‘iCAN!v0’ tool used for validation. Mean scores (MS; %) obtained from its administration to Greek medical graduates (n=508) and interpretation of the scores in color (see note 3)

Tuning Learning Outcomes (desired competences) ¹	The iCAN!v0 questionnaire ²	MS ³
Graduates in medicine will have the ability to:	Total Questionnaire	73
1 Carry out a consultation with a patient		76
1.1 take a history	65 I can take a correct history.	85
1.2 carry out physical examination	1 I can carry out a systematic physical examination.	84
1.3 make clinical judgments and decisions	94 I can make clinical judgments and take clinical decisions.	67
1.4 provide explanation and advice	12 I can provide explanation and advice to the patient.	76
1.5 provide reassurance and support	20 I can provide reassurance and support to the patient.	73
1.6 assess the patient's mental state	87 I can assess a patient's mental state.	71
2 Assess clinical presentations, order investigations, make differential diagnoses, negotiate a management plan		69
2.1 recognize and assess the severity of clinical presentations	33 I can recognize and assess the severity of clinical presentations.	71
2.2 order appropriate investigations and interpret the results ⁴	61 I can order appropriate investigations.	75
	77 I can interpret the results of the investigations.	76
2.3 make differential diagnoses	70 I can perform differential diagnosis.	74
2.4 negotiate an appropriate management plan with patients and carers	4 I can negotiate an appropriate management plan with patients and carers.	74
2.5 provide care of the dying and their families	39 I can provide care of the dying and their families.	62
2.6 manage chronic illness	8 I can manage chronic illness.	64
2.7 be aware of additional diagnostic and therapeutic options available within other healthcare professions (ECC p13) ⁴	18 I am aware of the additional diagnostic and therapeutic options available within other healthcare professions.	59
2.8 uphold the STEEEP principle of patient care (ECC p10) ⁴	19 I can keep the STEEEP patient care principle (safe, timely, effective, equitable, efficient, patient-centered).	63
3 Provide immediate care of medical emergencies, including first aid and resuscitation		70
3.1 recognize and assess acute medical emergencies	96 I can recognize and assess acute medical emergencies.	73
3.2 treat acute medical emergencies	101 I can treat acute medical emergencies.	65
3.3 provide basic First Aid	31 I can provide basic First Aid.	76
3.4 provide basic life support and cardio-pulmonary resuscitation according to current European guidelines	63 I can provide basic life support according to current international guidelines.	72
3.5 provide advanced life support according to current European guidelines ⁴	–	
3.6 provide trauma care according to current European guidelines	71 I can provide trauma care according to current international guidelines.	66
4 Prescribe drugs		64
4.1 prescribe clearly and accurately	58 I can prescribe clearly and accurately.	69
4.2 match appropriate drugs and other therapies to the clinical context	5 I can match appropriate drugs and other appropriate therapies to the clinical context.	58
4.3 review the appropriateness of drug and other therapies and evaluate potential benefits and risks	22 I can review the appropriateness of drug and other therapies and evaluate potential benefits and risks.	62
4.4 treat pain and distress	30 I can treat pain and distress.	69
5 Carry out practical procedures		73
5.1 measure blood pressure	49 I can measure blood pressure.	95
5.2 venepuncture	16 I can take venous blood.	89
5.3 cannulation of veins	9 I can perform cannulation of veins.	69
5.4 administer IV therapy and use infusion devices	40 I can administer IV therapy and use infusion devices.	55
5.5 subcutaneous and intramuscular injection ⁴	69 I can give a subcutaneous injection.	80
	55 I can give an intramuscular injection.	83
5.6 administer oxygen	88 I can administer oxygen.	75
5.7 move and handle patients	11 I can move and handle patients.	70

5.8 suturing	10 I can suture.	72
5.9 blood transfusion	83 I can perform blood transfusion.	39
5.10 bladder catheterization	86 I can perform bladder catheterization.	69
5.11 urinalysis	103 I can perform urinalysis.	62
5.12 electrocardiography	35 I can perform electrocardiography.	91
5.13 basic respiratory function tests	32 I can perform basic respiratory function tests.	71
6 Communicate effectively in a medical context		75
6.1 communicate with patients	48 I can communicate with patients.	84
6.2 communicate with colleagues	57 I can communicate with colleagues.	83
6.3 communicate in breaking bad news	85 I can communicate in breaking bad news.	63
6.4 communicate with relatives	54 I can communicate with patients' relatives.	79
6.5 communicate with disabled people	99 I can communicate with disabled people.	72
6.6 communicate in seeking informed consent	72 I can communicate in seeking informed consent.	72
6.7 communicate in writing (including medical records)	97 I can communicate in writing.	79
6.8 communicate in dealing with aggression	3 I can communicate in dealing with aggression.	73
6.9 communicate by telephone	102 I can communicate by telephone.	82
6.10 communicate with those who require an interpreter	29 I can communicate with those who require an interpreter.	67
6.11 be aware of their own and others' non-verbal communication (ECC p6) ⁴	23 I am aware of the non-verbal ways of communication used by others or myself.	70
7 Apply ethical principles in medical practice		70
7.1 maintain confidentiality	41 I can maintain confidentiality.	91
7.2 apply ethical principles and analysis to clinical care	89 I can apply moral and ethical principles to patient's care.	81
7.3 obtain and record informed consent	37 I can obtain and record informed consent.	72
7.4 certify death	93 I can complete correctly the medical certificate of death.	57
7.5 request autopsy	104 I know when and how to request autopsy.	51
7.6 apply national and European law to clinical care	42 I can apply national and European law related to patient's care.	69
8 Assess psychological and social aspects of a patient's illness		71
8.1 assess psychological factors in presentations and impact of illness	15 I can assess psychological factors in presentations and impact of illness.	70
8.2 assess social factors in presentations and impact of illness	25 I can assess social factors in presentations and impact of illness.	73
8.3 detect stress in relation to illness	75 I can detect stress caused by the illness.	70
8.4 detect alcohol and substance abuse, dependency ⁴	45 I can detect alcohol dependency	73
	82 I can detect substance dependency	70
9 Apply the principles, skills and knowledge of evidence-based medicine		68
9.1 apply evidence to practice	13 I can apply evidence to practice (evidence based medicine).	66
9.2 define and carry out an appropriate literature search	76 I can define and carry out an appropriate literature search.	71
9.3 critically appraise published medical literature	6 I can critically appraise published medical literature.	66
10 Use information and information technology effectively in a medical context		77
10.1 keep accurate and complete clinical records	81 I can keep accurate and complete clinical records.	76
10.2 use computers	21 I can use information systems used in healthcare.	70
10.3 access information sources	91 I can have access to information sources.	80
10.4 store and retrieve information	66 I can store and retrieve information.	81
11 Apply scientific principles, methods and knowledge to medical practice and research		73
<i>No specific level 2 outcomes⁴</i>	90 I can apply scientific principles, methods and knowledge in medical practice.	76
	100 I can apply scientific principles, methods and knowledge to medical research.	70
12 Promote health, engage with population health issues and work effectively in a health care system		73
12.1 provide patient care which minimizes the risk of harm to patients	2 I can provide to patient care which minimizes the risk of harm them.	73
12.2 apply measures to prevent the spread of infection	14 I can apply measures to prevent the spread of infection.	62

<i>12.3 recognize own health needs and ensure own health does not interfere with professional responsibilities⁴</i>	27 I can recognize my own health needs.	81
	44 I can ensure own health does not interfere with professional responsibilities.	79
12.4 conform with professional regulation and certification to practice	26 I can conform with professional regulation and certification to practice.	82
12.5 receive and provide professional appraisal	38 I can receive and provide professional appraisal.	66
12.6 make informed career choices	60 I can make informed career choices.	69
12.7 engage in health promotion at individual and population levels	79 I can engage in health promotion at individual and population level.	70
13 Professional attributes		82
13.1 probity, honesty, ethical commitment	56 I can commit by probity, honesty and morality & ethics.	88
13.2 commitment to maintaining good practice, concern for quality	28 I can devote to maintaining good practice and concerning for quality.	85
13.3 critical and self-critical abilities, reflective practice	47 I have critical and self-critical abilities.	82
13.4 empathy	51 I can empathize.	77
13.5 creativity	24 I can be creative.	80
<i>13.6 initiative, will to succeed⁴</i>	36 I can take initiatives.	73
	7 I have the will to succeed.	91
13.7 interpersonal skills	84 I can develop interpersonal relations.	79
14 Professional working		75
14.1 ability to recognize limits and ask for help	17 I can recognize my limits and ask for help.	87
14.2 capacity to deal with uncertainty and adapt to new situations	98 I can deal with uncertainty and adapt to new situations.	71
14.3 ability to lead others	78 I have capacity of leadership (ability to lead others).	67
14.4 ability to work autonomously when necessary	64 I can work autonomously when necessary.	72
14.5 ability to solve problems	68 I can solve problems.	73
14.6 ability to make decisions	95 I can make decisions.	74
14.7 ability to work in a multidisciplinary team	52 I can work in a interdisciplinary or interprofessional team.	76
14.8 ability to communicate with experts in other disciplines	43 I can communicate with experts in other disciplines.	79
<i>14.9 capacity for organization and planning (including time management)⁴</i>	62 I can organize and plan.	75
	46 I can manage my time.	75
15 The doctor as expert		72
15.1 capacity for analysis and synthesis	80 I have the capacity for analysis and synthesis.	75
15.2 capacity to learn (including lifelong self-directed learning)	53 I learned to learn in this school	71
15.3 capacity for applying knowledge in practice	74 I can apply my knowledge in practice.	74
15.4 ability to teach others	50 I can teach others.	71
15.5 research skills	67 I have research skills.	68
16 The global doctor		80
<i>16.1 appreciation of diversity and multicultural⁴</i>	92 I can appreciate diversity and multiculturality.	84
<i>16.2 understanding of cultures and customs of other countries⁴</i>	–	–
16.3 ability to work in an international context	34 I can work in an international context.	65
16.4 knowledge of a second language	73 I have knowledge of a second language.	88
16.5 general knowledge outside medicine	59 I have general knowledge outside medicine.	83
–	<i>105 I owe my medical knowledge and abilities to my school's curriculum.⁵</i>	61

¹ Outcome identification numbers were assigned consecutively to *The Tuning Learning Outcomes* (Cumming & Ross, 2008: pages 13-17) [2].

² The numbers in front of the questions are their identification numbers, assigned randomly using the Microsoft Excel function "rand()". Questions in the distributed questionnaire were sorted according to this number.

³ MS = standard mean score in the scale 0 to 100 (%), based on the tested version 'iCAN!v0'. The color of cells serves the 'at a glance' interpretation: red: very poor (score<50), respondents should not be graduated (pass/fail cut-point 50%); orange: poor (50-59); yellow: rather poor (60-69); gray: rather good (70-79); blue: good (80-89); green: very good (90-100); (MS in outcomes 5.7, 6.11, 10.2, 11b, and 12.7 are between 69.5 and 69.9 rounded to 70 but their color is correctly yellow). The last question (105) is not an outcome question; it refers to the educational process during the total study years and thus its interpretation and color is different than all other scores: red: very poor (score<25); orange: poor (25-39); yellow: rather poor (40-49); gray: rather good (50-59); blue: good (60-74); green: very good (75-100). ATTENTION: this is a validation study, not a study of the educational outcomes of Greek medical schools (see Limitations); the table and scores are given to help readers to understand the whole process of development and validation and not to describe Greece's medical education outcomes.

⁴ Italics denote a) the seven bipartite outcomes (six 'level two' and one 'level one') that were split into two separate questions, i.e., outcome 2.2 (that was split in the questions 61 and 77), 5.5 (55, 69), 8.4 (45, 82), 11 (90, 100), 12.3 (27, 44), 13.6 (7, 36) and 14.9 (46, 62); b) the two outcomes 16.1 and 16.2 that were

considered similar and combined in one question (92), c) the outcome that was considered inappropriate for undergraduate level and was omitted (3.5), and d) the three outcomes that were added to the questionnaire from the European Core Curriculum (ECC) and assigned the outcome identification numbers 2.7, 2.8, and 6.11.

⁵ This is not an outcome question, this is an educational process question raised during development (see note 3).

2.3. Testing the 'iCAN!' Tool with Graduates

The 'iCAN!v0' was first transformed by a private agency (Anova Consulting, www.anova.gr) to an anonymous scanable form (available upon request). Then a survey among Greek medical graduates was carried out, approved by the Deans of all Greek medical schools (see Acknowledgements). The scanable forms were distributed to a convenient sample in all seven medical schools, during the 2009-2011 graduation periods. Completed forms were scanned by the Anova Consulting, using a high reliability optical mark recognition scanner (OpScan iNSIGHT™, Pearson NCS) and the QuickTesting software. Thus, an electronic data file was obtained, which was checked against the original completed questionnaires by both the Anova Consulting and the authors.

According to SF-36 scoring instructions [48], response options were assigned scores in the six equal interval Likert scale 0-100 as follows: strongly disagree = 0, disagree = 20, rather disagree = 40, rather agree = 60, agree = 80, strongly agree = 100, and mean scores were calculated (instead of sums as in DREEM [33] and PHEEM [36]). This 'standard scoring method' [46] has the advantage that all absolute scores (overall, subscale or single question ones) coincide with their corresponding percentages, since both range from 0 to 100, the easiest scale for everybody to remember, interpret and compare.

Since the Tuning learning outcomes were set and described in two granularity levels ('level one' and 'level two') by consensus of a large qualitative group of teaching staff from over ninety European schools through the whole MEDINE process rather than by statistical methods [2], we considered more appropriate to base our analysis on 'level one' in order to test the properties of each Tuning 'level one' learning outcome and to identify whether their 'level two' learning outcomes compose a single factor (domain) or can be best described with a two or more factors model.

Cronbach's alpha and 'alpha if item deleted' were calculated to check internal consistency within each of the sixteen Tuning 'level one' learning outcomes. Cronbach's alpha was also calculated for various splits of the total questionnaire, i.e., by type of outcome (medical ability, professionalism), medical school (for each of the seven Greek medical schools), gender (male, female, not declared), and time to complete the questionnaire (the fast and the slow ones: the less and the more than the mean time to complete). The Spearman-Brown formula was used to estimate expected split alphas [45-46] to compare with the directly calculated ones.

Exploratory factor analysis was used to analyze interrelations among the questions within each of the sixteen Tuning 'level one' learning outcomes (construct validity). Seven questions (13, 16, 27, 35, 41, 49, 73) were transformed

before performing factor analysis since they did not satisfy normality [47] checked by skewness and kurtosis. The extraction method was Maximum Likelihood, and the rotation method Varimax with Kaiser Normalization. Factors accepted were those fulfilling the Schönrock-Adema et al [47] principles: scree plot point of inflexion; eigenvalues >1.5; proportion of variance accounted for an approximate additional 5% of the variance; a given factor should contain at least three variables with significant (>0.40) loadings; variables loading on the same factor share the same conceptual meaning; variables loading on different components appear to measure different constructs.

In order to test the instrument's ability to detect really existing differences among different groups, the total questionnaire mean scores among undergraduates, graduates, and practitioners were compared. The underlying hypothesis was that the higher the medical maturity is the higher the overall mean score must be. For this reason, in addition to graduates, seven groups of 1st to 6th year undergraduate medical students at the Ioannina University and a group of practicing physicians three to thirty years after graduation (median=10, Q₁=5.5, Q₃=15.5) were also included. ANOVA was first used to check whether mean scores differ, and then unpaired two-tailed t-test with unequal variances was used to locate the differences between the groups. Parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions, without the fear to come to the wrong conclusion [66].

Analyses were performed using Microsoft Excel 2003 and SPSS 16.0. Reported p-values were considered significant at p<0.05. Responsiveness (changes within a group over time) was beyond the scope of this study.

2.4. 'Ican!V1': The Released Version of the Questionnaire

Questions that seemed to be problematic in the previous testing phase, i.e. questions that if omitted lead to increased alphas or questions loaded less than 0.40 in a factor or loaded to two factors, were thoroughly re-examined. In addition, through the experience of the whole testing phase, we found many opportunities for a more penetrating glance at all questions enabling further refinement. Thus, in order to avoid misunderstandings and to be as clear as possible, we accepted periphrastic wording, including even definitions or glossies. For example, "I can take an appropriate history, detailed and relevant" instead of simply "I can take a history", or "I can make evidence based medical decisions, unbiased by any conflict of interest" instead of "I can make decisions": everybody makes decisions all the time, but we are not interested in graduate's personal life, only about medical decisions, more precisely whether they are

evidence-based and unbiased from any conflict of interest, direct or indirect, obvious or hidden; this is what we are really interested in, and in this way the participant is clear of what he is being asked.

Under the umbrella rule not to exclude any Tuning 'level two' learning outcome, outcome 3.5 (advanced life support) and 16.2 (understanding other countries), not included in the tested version (v0), were re-incorporated to the released version (v1): It is MEDINE's responsibility to eliminate Tuning learning outcomes. Since outcome 3.5 covers entirely outcome 3.4 (basic life support), the corresponding to outcome 3.4 question could be omitted. However, since some graduates might be able in outcome 3.4 but not in 3.5, it was decided to keep both. Outcomes 3.4–3.6 ask furthermore if the graduates comply with the European guidelines. However, not the legislation, European, American, Greek etc., is here meant, but the evidence-based and scientific guidelines; thus we replaced "European" with "scientific (evidence-based)".

Tuning 'level two' learning outcome 1.3 states "Graduates in medicine will have the ability to make clinical judgments and decisions". However, everybody makes judgments and decisions, but making the right clinical judgment and evidence-based decision is what we actually are interested in when asking the medical graduate. These two words make all the difference and the real underlying meaning of this outcome is that the "right person does the right thing right" [6]. Therefore we rephrased the statement "I can make clinical judgments and take clinical decisions" to "I can make right clinical judgments and evidence-based medical decisions" in the released version.

The Tuning 'level one' learning outcome 6 states "Graduates in medicine will have the ability to communicate effectively in a medical context". The critical word here is "effectively", but it was not included in version v0. We had to insert it in all corresponding questions in the released version.

Regarding the Tuning 'level two' learning outcome 1.1 "to take a history", we are interested in to learn from the medical graduate whether he has the ability to take an appropriate history, as formulated in the Mexican learning outcome ("The ability to take an appropriate history") [15], since everyone (any healthcare professional, the patient, the relatives, even the lay people) can "take a history". Thus, we formulated the correspondent statement as "I can take an appropriate history, detailed and relevant" in the released version 'iCAN/v1'.

Two additional questions ("learn from own mistakes" and "bear in mind the consequences of own decisions") from the European Core Curriculum [26] were considered important professionalism attributes, especially the first one [54-60] and that is why they were included in 'iCAN/v1'. The same applies for the last but one question, aiming at the overall self-assessment of the graduate.

Finally, based on graduates declaration attributing their knowledge and abilities to medical school's curriculum by about 60% [61] we added the first open question for the

remaining 40% to clarify where they attribute their knowledge and abilities if not to the curriculum.

Furthermore we realized that the open question "If you could change one thing in your school, what would this be?" was more suitable for an educational environment metric rather than for an educational outcome metric [43] and thus we replaced it with two open questions "What medical competence, which you are not competent at, you would like to be competent?" and "what medical competence, which you are competent at, you would not bother if you were not competent at?" Both questions could also serve for future content revisions of medical learning outcomes.

3. Results

3.1. Graduates Participating in 'iCAN/v0' Testing and Data

Out of 899 graduates of twelve graduating cohorts between July 2009 and November 2011 from all seven Greek medical schools, 512 participated (response rate 57%); 43% male, 57% female; 77% graduated in July 2009, 23% afterwards; 148 out of 310 from Athens University (48%), 32/68 (47%) from Crete, 12/72 (17%) from Ioannina, 61/120 (51%) from Patras, 147/208 (71%) from Thessaloniki, 74/74 (100%) from Thessaly, and 38/47 (81%) from Thrace.

It took them 15 to 20 minutes to complete the questionnaire (mean 18, standard deviation 10; median 15, interquartile range 10 to 20).

Three of the 512 questionnaires were empty except the open question that was answered, and one questionnaire was answered with the option 'strongly agree' in all closed questions. These four questionnaires were excluded. Six more questionnaires had overlooked page 3 for technical reasons (the unusual top-right stitching in order to keep clear the left edge of the form for the scanner to work correctly), resulting in 174 missing values. In addition, there were 371 missing values (0.7%, ranged 0% in q1 to 1.6% in q100) scattered throughout the dataset without any obvious underlying pattern. For factor analysis, these 545 (174+371) missing values, 1.03% of the whole dataset, much less than 20% [67], were replaced with the variable mean [47]. The 52,391 non-missing values were distributed to strongly agree 25%, agree 36%, rather agree 26%, rather disagree 9%, disagree 3%, strongly disagree 1%. In only 8 of the 104 questions (Table 1), the 'strongly agree' option counted for more than 50% of all six options: q49 'measure blood pressure' 83%, q35 'electrocardiography' 69%, q7 'will to succeed' 65%, q41 'confidentiality' 64%, q16 'venepuncture' 59%, q73 'knowledge of a second language' 59%, q56 'probity, honesty, ethical commitment' 54%, and q17 'ability to recognize own limits and ask for help' 51%.

3.2. Internal Consistency

Cronbach's alphas are given in Table 2. In the first scale,

for example, "Consultation with a patient", consisting of six questions, from the 508 questionnaires 491 had no missing values, giving an acceptable alpha of .765. Five of the sixteen scale alphas were good ($>.8$), eight acceptable ($>.7$), and only three questionable ($>.6$). 'Alpha if item deleted' was slightly higher in five scales (5, 7, 9, 14 and 15), in the second decimal place and beyond, that could reasonably be considered negligible (see notes a-e). Alphas by type of outcome, medical school, gender, and time to complete the questionnaire were $>.9$, as expected by the Spearman - Brown formula, in part due to the high number of questions.

3.3. Factor Analysis

All corrected item total correlations were above .200 (between .214 and .706). Thus, all questions were kept for the factor analysis, a synopsis of which is presented in Table 3. For the first scale for example, "Consultation with a patient", a medical ability (m) Tuning 'level one' learning outcome (TL1LO), all its six questions (12, 94, ..., 1) were loaded to only one factor with loadings from .719 maximum (q12) to .636 minimum (q1), eigenvalue 2.8, percent of variance explained 45.9, and percentage of the 'Consultation with a patient' explained by these six questions 98.9%. In summary, fifteen of the sixteen TL1LO were explained by only one factor each, and only the fifth was split in two factors. The fifteen factors accounted for 39% to 80% of the variance of each TL1LO, with eigenvalues between 1.6 and 4.4, and loadings between .401 and .893, with the only exception for the question 17 that was slightly lower, .344 (see note 2 in Table 3): This solution was factorially simple (all questions of each TL1LO were loaded on one factor) and interpretable (the same meaning as the original TL1LO). The fifth TL1LO was split in two factors, with eleven questions the first factor (eigenvalue 4.3, variance explained 35.3%, loadings between .470 and .745), and five the second (eigenvalue 2.4, variance explained 13.3%, loadings between .455 and .761), two of which were loaded on both ($>.400$): This solution is also acceptable; the second factor contains the 4 or 5 most common or simple or taught or known practical procedures, and the first the remaining 10 or 9 less common/ simple/ taught/ known ones.

In general, all factors met the Schönrock-Adema et al [47] criteria described in Methods, with few marginal exceptions: Only one of the seventeenth factors, the 11th, contained less than three questions, due to that the eleventh TL1LO contained only two questions (q100, q90). Only one of the 104 questions loaded slightly lower than .4 (q17 "Ability to recognize own limits and ask for help", .344). Two

questions (q55 "Intramuscular injection", rather common; q69 "Subcutaneous injections", rather un-common) loaded $>.4$ in both the fifth TL1LO factors, which, though, could be interpreted as the less and the more common practical procedures (Table 3). At any rate, these questions should be considered carefully in the next version of the questionnaire.

3.4. Sensitivity (Discriminant Validity)

The total questionnaire mean score (TMS) increased from 42.6% in first year students to 81.4% in practitioners (Table 4). The increment from group to group was statistically significant ($p<0.001$), with the only exception of the two fifth-year student groups where, as expected, no significant difference was observed. The opposite occurred for standard deviation (not shown) that decreased from 33.3 to 20.4. That is, as participants mature, they feel more competent (higher scores) and more self-confident (lower deviation). These results are in line with what could be expected.

Among graduates (Table 1), the scale mean score varied from 64% (prescribing drugs) to 82% (professional attributes), while the question mean score ranged from 39% (blood transfusion) to 95% (measuring blood pressure). It is not known whether scale mean score is in line with what it is expected, but it is very plausible that question mean score is as declared, since every student should be able to measure blood pressure, but no Greek medical curriculum might include blood transfusion in their outcomes. This is an indication that the questionnaire is able to differentiate well between the weak and strong areas.

3.5. Final Refinement of the Tool: from the Tested 'Ican!V0' to the Released 'Ican!VI'

Questions that seemed to be problematic in the previous testing phase, i.e., questions that if deleted increased alphas even though by a negligible amount (from .006 to .052 if questions 49, 35, 41, 17, 13 and 53 were deleted) or loaded less than .40 in one factor (only q17) or loaded to two factors (q55, q69), were thoroughly re-examined. In addition, going through the whole testing phase, we experienced many opportunities for a more penetrating glance at all questions that enabled further refinement.

After the validation experience and questionnaire's refinement as described in last subsection of the Material and Methods chapter (2.4), Table 5 presents the hereby released version 'iCAN!VI'.

Table 2. Cronbach's alpha by Tuning 'level one' learning outcomes, type of outcome, medical school, gender, completeness time of the questionnaire and overall tool

	Question	Participants	Cases	Alpha	Interpretation
Scale (level one tuning learning outcome)					
1 Consultation with a patient	6	508	491	.765	Acceptable
2 Assess clinical presentations	9	508	484	.849	Good
3 Emergencies immediate care	5	508	494	.850	Good
4 Prescribe drugs	4	508	491	.734	Acceptable
5 Carry out practical procedures	14	508	481	.851a	Good
6 Communicate in medical context	11	508	485	.832	Good
7 Apply ethical & legal principles	6	508	496	.709b	Acceptable
8 Assess illness psycho-social aspects	5	508	491	.773	Acceptable
9 Apply evidence-based medicine	3	508	490	.736c	Acceptable
10 Use information & IT in medicine	4	508	494	.674	Questionable
11 Apply science to practice & research	2	508	499	.744	Acceptable
12 Healthcare system Population health	8	508	488	.783	Acceptable
13 Professional attributes	8	508	485	.789	Acceptable
14 Professional working	10	508	491	.852d	Good
15 The doctor as expert	5	508	490	.691e	Questionable
16 The global doctor	4	508	495	.612	Questionable
Type of outcome					
Medical abilities (scales 1 to 12)	77	508	429	.971	Excellent
Professionalism (scales 13 to 16)	27	508	472	.923	Excellent
School (all scales)					
Athens	104	148	126	.971	Excellent
Thrace	104	38	28	.959	Excellent
Thessaloniki	104	146	119	.979	Excellent
Ioannina	104	12	8	.982	Excellent
Crete	104	32	25	.978	Excellent
Thessaly	104	73	62	.986	Excellent
Patras	104	59	46	.977	Excellent
Gender (all scales)					
Male	104	159	130	.975	Excellent
Female	104	209	172	.982	Excellent
Not declared	104	140	112	.974	Excellent
Time to complete (all scales)					
< mean	104	237	192	.977	Excellent
≥ mean	104	155	133	.980	Excellent
Not declared	104	116	89	.976	Excellent
Total questionnaire (all scales)	104	508	414	.978	Excellent

Questions: the number of questions per scale.

Participants: the number of participants, i.e., the number of suitable completed questionnaires (508 in total, after excluding the three ones that had not answered any closed question and one that had answered 'strongly agree' to all questions).

Cases: the number of questionnaires without any missing value, as reported by SPSS; only these were used in alpha calculations.

Alpha: observed Cronbach's alpha.

Interpretation of the observed alphas: Excellent: >.9, Good: >.8, Acceptable: >.7, Questionable: >.6, P = poor (≥.5), Unacceptable: <.5.

a increased by .003 to .854 if question 49 or 35 were deleted.

b increased by .006 to .715 if question 41 was deleted.

c increased by .011 to .747 if question 13 was deleted.

d increased by .007 to .859 if question 17 was deleted.

e increased by .052 to .743 if question 53 was deleted.

Table 3. The original Tuning 'level one' learning outcomes (TL1LO) and synopsis of the exploratory factor analysis (EFA) of the 'iCAN!v0' data (n= 508 graduates)

Tuning 'level one' learning outcomes (TL1LO)				The factors revealed by the exploratory factor analysis (EFA 1)							
ID	Definition	T	Q	F	S	EV	%V	Q/F	Questions in loading descending order ²	Loading max, min	LRA adj R ²
1	Consultation with a patient	m	6	1	1	2.8	45.9	6	12, 94, 20, 65, 87, 1	.719, .636	.989
2	Assess clinical presentations	m	9	1	1	4.2	46.3	9	33, 61, 77, 70, 4, 39, 8, 18, 19	.751, .591	.979
3	Emergencies immediate care	m	5	1	1	3.1	62.7	5	101, 71, 31, 63, 96	.820, .753	.990
4	Prescribe drugs	m	4	1	1	2.2	56.1	4	5, 22, 30, 58	.797, .671	.989
5	Carry out practical procedures	m	14	2	2	4.3	35.3	11	40, 9, 88, 10, 83, 86, 32, 103, 69, 11, 55	.745, .470	.958
						2.4	13.3	5	35, 49, 55, 16, 69	.761, .455	
6	Communicate in medical context	m	11	1	1	4.3	39.2	11	48, 54, 72, 102, 99, 57, 97, 29, 3, 85, 23	.733, .471	.978
7	Apply ethical & legal principles	m	6	1	1	2.5	41.2	6	93, 37, 42, 104, 89, 41	.708, .439	.978
8	Assess illness psychosocial aspects	m	5	1	1	2.6	52.5	5	25, 75, 82, 15, 45	.756, .705	.979
9	Apply evidence-based medicine	m	3	1	1	2.0	65.2	3	6, 76, 13	.842, .733	.987
10	Use information & IT in medicine	m	4	1	1	2.1	51.5	4	66, 91, 81, 21	.751, .643	.991
11	Apply science to practice & research	m	2	1	1	1.6	79.7	2	100, 90	.893, .893	.995
12	Work in healthcare system; Population health	m	8	1	1	3.2	39.8	8	38, 79, 44, 26, 14, 60, 2, 27	.693, .553	.986
13	Professional attributes	p	8	1	1	3.3	41.0	8	47, 56, 24, 28, 84, 7, 51, 36	.709, .562	.970
14	Professional working	p	10	1	1	4.4	43.6	10	62, 68, 95, 64, 98, 78, 43, 52, 46, (17)	.788, .595 (.344)	.988
15	The doctor as expert	p	5	1	1	2.4	47.7	5	80, 67, 74, 50, 53	.774, .401	.990
16	The global doctor	p	4	1	1	1.9	46.6	4	73, 59, 34, 92	.751, .624	.979

¹ Rotated factor matrix, factors 17 (all those with eigenvalues greater than 1.5, and loading suppress .4); Extraction method: Maximum Likelihood; Rotation method TL1LO 5: Varimax with Kaiser Normalization; Rotation converged in less than 25 iterations; Seven items (13, 16, 27, 35, 41, 49, 73) were transformed before performing factor analysis because they did not satisfy normality, checked by skewness and kurtosis.

² Questions' ID (as in Table 1) in descending order by question loading: the first question had the maximum loading, and the last one the minimum. 106 questions in total, because q55 and q69 are presented twice, in bold and plain (higher and lower loadings, respectively), due to their loading in both sub-factors of the fifth factor. The TL1LO 14 shows in parentheses the only question (17) with loading <.4 (.344), while the other nine questions of this factor had loadings between .788 (q62) and .595 (q46).

ID: the TL1LO identification number as in Table 1.

T: type of the outcome (m = medical ability; p = professionalism; 16 in total, 12m+4p).

Q: number of questions per outcome (104 in total).

F: number of factors per TL1LO (17 in total, 16+1, due to the fifth TL1LO splitting in two factors).

S: scree plot sharp point of inflection after the ith factor (i=2 for the fifth TL1LO, i=1 for all the rest).

EV: eigenvalue (between 1.6 and 4.4).

%V: percent of variance explained.

Q/F: questions per factor (equal to Q, except in TL1LO 5, where two questions were loaded to both factors, see note 2 above).

Loading (max, min): loading range from maximum to minimum of the questions of the factor (presented in the previous column: the max loading corresponds to the first question and the min to the last, while the loadings of all other questions of the factor lie in between).

LRA: linear regression analysis.

Adj R2: adjusted R square = the percentage of the TL1LO explained by all its questions that are statistically significantly related with it (all question are highly significantly related with corresponding TL1LO, p<.001).

Table 4. Discriminant validity: Total questionnaire mean scores by various medical maturity groups.

GROUP	P	N	TMS (95% CI)	t	p
1 st year students	73	7547	42.6 (41.8–43.3)		
2 nd " "	51	5293	44.4 (43.6–45.3)	3.3	0.001
3 rd " "	112	11496	47.7 (47.2–48.3)	6.3	<0.001
4 th " "	39	4005	52.2 (51.3–53.0)	8.5	<0.001
5 th " "1	36	3684	57.2 (56.3–58.1)	8.2	<0.001
5 th " "2	14	1455	57.7 (56.4–59.0)	0.6	0.527
6 th " "	18	1861	61.2 (60.1–62.3)	4.1	<0.001
Graduates	508	52287	73.3 (73.2–73.6)	21	<0.001
Practitioners	21	2119	81.4 (80.5–82.3)	17	<0.001

P: participants; N: number of options chosen by the participants among the six offered ones (from strongly disagree to strongly agree); TMS: total questionnaire mean scores; CI: confidence interval.

t: unpaired two-tailed t-test with unequal variances (if equal, a rather reasonable condition, p-values, not shown, are negligibly smaller). ANOVA was previously performed (p=0.000)

p: p-value comparing the corresponding mean with the one immediately above it; e.g., the first p compares the second TMS with the first one, the second p compares the third TMS with the second one etc.; in bold the only non significant.

1, 2: different fifth year students from two successional cohorts, 20101 and 20092.

Table 5. The 'iCAN! v1' questionnaire ready to use by a medical school to obtain its graduates' perceptions on their own medical abilities and professionalism

Dear graduate, congratulations! Have a great career!						
Your medical school awarded you your bachelor degree. Now it is your turn: you can help your school, coming generations of colleagues, and tomorrow's patients and healthy. Do you agree with the following statements? To what extent do they represent you? This is about how you personally feel at this moment.						
	SD strongly disagree	D disagree	RD rather disagree	RA rather agree	A agree	SA strongly agree
1. I can carry out correctly a systematic physical examination, both general and targeted	SD	D	RD	RA	A	SA
2. I can provide effectively patient care that minimises the risk of harm to patient	SD	D	RD	RA	A	SA
3. I can communicate effectively in dealing with aggression	SD	D	RD	RA	A	SA
4. I can negotiate effectively with the patient and carers an appropriate management plan of their illness	SD	D	RD	RA	A	SA
5. I can correctly match appropriate drugs and other appropriate therapies to the clinical context	SD	D	RD	RA	A	SA
6. I can critically appraise published medical literature	SD	D	RD	RA	A	SA
7. I can achieve to be the good tomorrow's doctor	SD	D	RD	RA	A	SA
8. I can manage effectively chronic illness	SD	D	RD	RA	A	SA
9. I can properly perform cannulation of veins	SD	D	RD	RA	A	SA
10. I can properly suture	SD	D	RD	RA	A	SA
11. I can move and handle patients without being at risk, neither the patients nor me	SD	D	RD	RA	A	SA
12. I can provide effectively explanation and advice to the patient	SD	D	RD	RA	A	SA
13. I can effectively apply to medical practice the best evidence available (evidence-based medicine)	SD	D	RD	RA	A	SA
14. I can apply effectively measures to prevent the spread of infection	SD	D	RD	RA	A	SA
15. I can assess correctly the psychological factors in presentations and impact of illness	SD	D	RD	RA	A	SA
16. I can properly venepuncture	SD	D	RD	RA	A	SA
17. I can recognize effectively my own limits and ask for help	SD	D	RD	RA	A	SA
18. I can use effectively the additional diagnostic and therapeutic options available within other healthcare professions	SD	D	RD	RA	A	SA
19. I can uphold undeviatingly the patient care principle STEEEP (safe, timely, effective, equitable, efficient, patient-centred)	SD	D	RD	RA	A	SA
20. I can provide effectively reassurance and support to the patient	SD	D	RD	RA	A	SA
21. I can use effectively information systems used in healthcare	SD	D	RD	RA	A	SA
22. I can review effectively the appropriateness of drug and other therapies and evaluate potential benefits and risks	SD	D	RD	RA	A	SA
23. I can communicate effectively in non-verbal (body) language	SD	D	RD	RA	A	SA
24. I can be creative while practicing	SD	D	RD	RA	A	SA
25. I can assess correctly the social factors in presentations and impact of illness	SD	D	RD	RA	A	SA
26. I can conform appropriately to professional regulation and certification to practise	SD	D	RD	RA	A	SA
27. I can recognise properly my own health needs	SD	D	RD	RA	A	SA
28. I can commit undeviatingly to maintain good practice and to concern for quality of care and of life	SD	D	RD	RA	A	SA
29. I can communicate effectively with those who require an interpreter	SD	D	RD	RA	A	SA
30. I can treat effectively pain and distress, with or without drugs as appropriate	SD	D	RD	RA	A	SA
31. I can provide effectively basic First Aid according to current scientific (evidence-based) guidelines	SD	D	RD	RA	A	SA
32. I can correctly perform basic respiratory function tests	SD	D	RD	RA	A	SA
33. I can recognise and assess correctly the severity of clinical presentations	SD	D	RD	RA	A	SA
34. I can work effectively in an international context	SD	D	RD	RA	A	SA
35. I can take a correct electrocardiograph	SD	D	RD	RA	A	SA
36. I can take the necessary initiatives while practicing medicine	SD	D	RD	RA	A	SA
37. I can obtain and record properly informed consent	SD	D	RD	RA	A	SA
38. I can receive and provide constructive professional appraisal	SD	D	RD	RA	A	SA

39. I can provide effectively care of the dying and their families	SD	D	RD	RA	A	SA
40. I can properly administer intravenous therapy and use infusion devices	SD	D	RD	RA	A	SA
41. I can maintain properly confidentiality	SD	D	RD	RA	A	SA
42. I can apply correctly national and European law to clinical care	SD	D	RD	RA	A	SA
43. I can communicate effectively with experts in other disciplines or professions	SD	D	RD	RA	A	SA
44. I can ensure effectively that own health does not interfere with my professional responsibilities	SD	D	RD	RA	A	SA
45. I can detect effectively alcohol abuse and dependency	SD	D	RD	RA	A	SA
46. I can manage effectively my own time	SD	D	RD	RA	A	SA
47. I can receive and provide constructive criticism and self-criticism	SD	D	RD	RA	A	SA
48. I can communicate effectively with patients	SD	D	RD	RA	A	SA
49. I can correctly measure blood pressure	SD	D	RD	RA	A	SA
50. I can effectively teach others	SD	D	RD	RA	A	SA
51. I can be empathetic (empathy)	SD	D	RD	RA	A	SA
52. I can work constructively in a disciplinary or multiprofessional team	SD	D	RD	RA	A	SA
53. I can learn effectively, including lifelong self-directed learning	SD	D	RD	RA	A	SA
54. I can communicate effectively with patients' relatives	SD	D	RD	RA	A	SA
55. I can give correctly intramuscular injection	SD	D	RD	RA	A	SA
56. I can commit uncompromisingly to probity, honesty, ethics and morality	SD	D	RD	RA	A	SA
57. I can communicate effectively with colleagues	SD	D	RD	RA	A	SA
58. I can prescribe clearly and accurately	SD	D	RD	RA	A	SA
59. I can use in my practice general knowledge outside medicine	SD	D	RD	RA	A	SA
60. I can make effectively informed career choices	SD	D	RD	RA	A	SA
61. I can order appropriate (evidence-based) investigations, those that are really needed	SD	D	RD	RA	A	SA
62. I can effectively organize and plan my medical work	SD	D	RD	RA	A	SA
63. I can provide effectively basic life support (BLS) and cardio-pulmonary resuscitation (CPR) according to current scientific (evidence-based) guidelines	SD	D	RD	RA	A	SA
64. I can effectively work autonomously as a physician when necessary	SD	D	RD	RA	A	SA
65. I can take an appropriate history, detailed and relevant	SD	D	RD	RA	A	SA
66. I can effectively store and retrieve medical information	SD	D	RD	RA	A	SA
67. I can work as a researcher (I have obtained the appropriate skills)	SD	D	RD	RA	A	SA
68. I can solve effectively health problems	SD	D	RD	RA	A	SA
69. I can give correctly subcutaneous injections	SD	D	RD	RA	A	SA
70. I can make correct differential diagnoses	SD	D	RD	RA	A	SA
71. I can provide effectively advanced trauma care (ATLS) according to current scientific (evidence-based) guidelines	SD	D	RD	RA	A	SA
72. I can communicate effectively in seeking informed consent	SD	D	RD	RA	A	SA
73. I can communicate effectively in a second language	SD	D	RD	RA	A	SA
74. I can correctly apply knowledge in practice	SD	D	RD	RA	A	SA
75. I can detect effectively stress attributed to illness (the illness as a cause of stress)	SD	D	RD	RA	A	SA
76. I can define and carry out effectively an appropriate literature search	SD	D	RD	RA	A	SA
77. I can interpret correctly the results of investigations, no matter if ordered by me or others	SD	D	RD	RA	A	SA
78. I can lead effectively a healthcare team	SD	D	RD	RA	A	SA
79. I can engage effectively in disease prevention and health promotion at individual and population levels	SD	D	RD	RA	A	SA
80. I can break down a health-related situation into its components (analysis) and combine separate elements to a coherent whole (synthesis)	SD	D	RD	RA	A	SA

81. I can keep effectively accurate and complete clinical records	SD D RD RA A SA
82. I can detect effectively substance abuse and dependency	SD D RD RA A SA
83. I can properly perform blood transfusion	SD D RD RA A SA
84. I can create effectively good interpersonal relations	SD D RD RA A SA
85. I can communicate effectively in breaking bad news	SD D RD RA A SA
86. I can properly perform bladder catheterisation	SD D RD RA A SA
87. I can assess correctly the patient's mental state	SD D RD RA A SA
88. I can properly administer oxygen	SD D RD RA A SA
89. I can apply correctly ethical principles and analysis to clinical care	SD D RD RA A SA
90. I can apply properly the principles, methods and knowledge of basic sciences to medical practice (in diagnosis, treatment, prognosis, and prevention)	SD D RD RA A SA
91. I can find and access effectively the appropriate information sources	SD D RD RA A SA
92. I can appreciate diversity and multiculturalism	SD D RD RA A SA
93. I can complete correctly the death certificate	SD D RD RA A SA
94. I can make right clinical judgements and evidence based clinical decisions	SD D RD RA A SA
95. I can make evidence-based medical decisions, unbiased by any conflict of interest	SD D RD RA A SA
96. I can correctly recognise and assess acute medical emergencies	SD D RD RA A SA
97. I can communicate effectively in writing	SD D RD RA A SA
98. I can deal effectively with uncertainty and adapt to new situations	SD D RD RA A SA
99. I can communicate effectively with disabled people	SD D RD RA A SA
100. I can apply properly the principles, methods and knowledge of basic sciences to medical research	SD D RD RA A SA
101. I can treat effectively acute medical emergencies	SD D RD RA A SA
102. I can communicate effectively by telephone	SD D RD RA A SA
103. I can correctly perform urinalysis	SD D RD RA A SA
104. I can request autopsy, and I know when and how to do it correctly	SD D RD RA A SA
105. I can learn effectively from my medical mistakes	SD D RD RA A SA
106. I can always bear in mind the consequences of my decisions and actions	SD D RD RA A SA
107. I can provide effectively basic trauma care (BTLS) according to current scientific (evidence-based) guidelines	SD D RD RA A SA
108. I can assess correctly the impact of stress and other psychological factors in patient's health (the stress as a cause of illness)	SD D RD RA A SA
109. I can diagnose correctly death	SD D RD RA A SA
110. I can provide effectively advanced life support (ALS) according to current scientific (evidence-based) guidelines	SD D RD RA A SA
111. I can commit to reflective, not to reflexive, practice	SD D RD RA A SA
112. I can understand cultures and customs of other countries	SD D RD RA A SA
113. I can assess correctly the impact of illness in patient's social condition (the illness as a cause of social problems)	SD D RD RA A SA
114. Overall, I am a very good doctor	SD D RD RA A SA
115. I am excited that I entered in medical school	SD D RD RA A SA
116. I owe my medical knowledge and abilities to my school's curriculum	SD D RD RA A SA

If you do not agree (A) or strongly agree (SA) in the last question, where do you attribute your knowledge and abilities? _____

What medical competence, which you are not competent at, you would like to be competent? Be as specific as possible: _____

And what medical competence, which you are competent at, you would not bother if you were not? Be as specific as possible: _____

Make sure you have answered all questions and you have given only one answer for each one.

Thank you for your collaboration!

4. Discussion

This study's goal was to describe the transformation of the Tuning learning outcomes to a useful questionnaire. Transforming a given outcome to a reasonable question was not an easy task as it first appeared. Thus, the discussion will focus to four interrelated but also independently important features of a questionnaire: reliability, validity, sensitivity, and responsiveness. Finally, the released version will be presented, followed by study's limitations, and the overall conclusion.

4.1. Validation

4.1.1. Internal Consistency and Reliability

All total questionnaire scale alphas were >0.95 , indicating excellent internal consistency [44, 46]. There are at least three reasons for such high alphas: the correlation of the items on the scale (actual reliability), the length of the scale (the number of questions), and the length of the Likert scale (the number of response options). Response options, five in educational environment metrics (DREEM, PHEEM etc), are six in 'iCAN!' and this adds to reliability [44]. The length of the total questionnaire (104 questions) is rather long; however, observed scale alphas were higher than or equal to expected ones, calculated by the Spearman - Brown formula that standardizes by the number of items of the total questionnaire [45, 46]. Supposing that other than actual reliability reasons cause some increase in Cronbach alpha, such as 15%, again total questionnaire overall alphas would remain good (>0.80); even accepting that much cutting as 25%, they would remain >0.70 , indicating an at least acceptable level of internal consistency. Thus, we can conclude that our efforts produced a consistent questionnaire.

4.1.2. Validity

Content validity and face validity both are optimized by a wide range of individuals involved [46]. The wide number of people involved in the Tuning-project, through literature searches, taskforce workshops and web-based opinion survey [2], insures content. The wide number of people involved in content's transformation into the 'iCAN!' questionnaire and in pilots insures that this content has been successfully transmitted into the questionnaire. The fact that all six alternative response options in almost all questions were used is in agreement with this conclusion [49]. Eleven questions with zero or only one answer in the 'strongly disagree' and 'disagree' options, as well as all questions with the majority of the answers in the rather uncertainty area of 'rather disagree' and 'rather agree', were thoroughly re-examined and appropriately rephrased in hereby released version 'iCAN!v1'.

Factor analysis plays a major role in construct validation [46]; in general, both confirmatory and exploratory factor analysis can be used due to study purposes. Because of the MEDINE consensus, we kept in the hereby released version

'iCAN!v1' the two items loaded in both factors in the fifth level one learning outcome and the one loaded slightly lower than 0.4 in the fourteenth 'level one' learning outcome (Table 3).

There is no other 'gold standard' or well-established instrument and this is not a prognostic study; thus, concurrent and prognostic validity remain unchecked.

4.1.3. Sensitivity and Responsiveness

High reliability is usually a prerequisite for sensitivity [46], and this condition was met as discussed before. Furthermore, the score differences found among undergraduates, graduates and practitioners are a good sign of the instrument's ability to detect existing differences among groups (sensitivity). It is reasonable to accept that these differences do really exist. It is also reasonable to accept that the two different groups of fifth year students should not differ (and 'iCAN!' found no difference), and that as medical participants mature their self-confidence increase (and iCAN! indicated exactly this: the standard deviation decreased). Another good indication of instrument's ability to detect existing abilities is the two extreme scores, the best (measuring blood pressure 95%; any medical student is expected to have been taught and learned this outcome) and the worst (blood transfusion 39%, far below the pass/fail cut-point of 50%; no undergraduate Greek curriculum contains this outcome). A rather unexpected result was the relatively high total mean score for the students of the first and second year of study. This could be at least partially attributed to the potential lower accurateness in the self-assessment the students may have in the first years of study compared to higher year students. Concluding, it is reasonable to accept that the tool has the sensitivity to discriminate strengths and weaknesses, offering thus a basis for a SWOT educational policy [62].

Testing the ability of the developed questionnaire to detect changes over time within the same group (responsiveness) was beyond the scope of this study. Thus, we have no indication of its responsiveness other than its sensitivity, since a highly sensitive scale will usually also be highly responsive [46]. It could also be reasonably argued that the changes being found between progressively matured groups (students, graduates, practitioners) carry a flavor of change over time, and this is in line with good responsiveness.

4.1.4. From 'iCAN!v0' to the Released Version of the Questionnaire 'iCAN!v1'

As discussed above, the validation exercise failed to reveal a non-reliable, non-valid, non-sensitive and perhaps non-responsive tool. However, a deeper insight in published literature and a more penetrating glance at all questions enabled further refinement, presented in Table 5. The equivalent Greek version is available from the corresponding author upon request.

4.2. Limitations

The translation of the original English Tuning learning outcomes to Greek language was not translated backwards [41]. However, 'iCAN!' creation is not a translation of an existing English questionnaire; it is a construction of the Greek 'ΜΙΤΟΡΩ!' simultaneously with its English counterpart 'iCAN!'. Since the original Tuning learning outcomes are in English and all authors have a good command of English and most of them have studied, lived and worked abroad for rather long periods of time and are publishing in English, we think that 'iCAN!' conveys successfully the meaning of the original outcomes.

All questions have been treated as equally weighted. However, outcomes 2.3 and 2.4 (differential diagnosis, mean score 74%, and treatment/management plan, 62%) are almost the whole medicine, with much more weight than measuring blood pressure (outcome 5.1, mean score 95%) or certifying death (outcome 7.4, mean score 57%). Weighting is a matter mainly of the Tuning project, and future research could answer whether question scores should be weighted.

Sometimes students and especially teachers considered Tuning 'level two' learning outcomes upon which the 'iCAN!' was built as too generic and not medically specific. They had a feeling that the biomedical paradigm that drives today's medicine and medical schools (anatomy, physiology, biochemistry, pathology etc) was not enough represented. Perhaps the knowledge outcomes described in Appendix A of 'The Tuning Project (Medicine)' [2] should be incorporated in the questionnaire. For example, "I can sufficiently to graduation level demonstrate knowledge of the normal function of all the systems of the human body (physiology)" or "...of the human society (sociology)" etc; however this would add about forty more questions. Perhaps also clinical attachments and experiential learning of Appendix B of 'The Tuning Project (Medicine)' should be included, e.g. "I had had sufficient to graduation level experienced clinical work in obstetrics and gynecological care"; beyond that they are process rather than outcomes, some fifteen more questions would be added. In the contrary, Tavakol & Dennick [64] argue that a value of Cronbach alpha >0.90 may suggest redundancies and show that the test length should be shortened.

Perhaps the main problem of the 'iCAN!' is its length. However it was somewhat compulsory, since the questionnaire maps one-to-one very closely to the Tuning 'level two' learning outcomes. Maybe a further refined version of these outcomes, a matter of the MEDINE network, could arise to a shortened version of the 'iCAN!' as well. On the other hand, one should keep in mind that the outcome of a medical school with six years of study might be difficult to be described by, say, thirty questions. Thus, further work is needed in these areas, we believe by the MEDINE network.

A warning rather than a limitation is that, although it was a national administration, the tool was not validated yet and our sample was not representative. Various sampling schemes, response rates, and cohort sizes across medical schools and time periods make the data appropriate for testing the instrument under construction than for detecting

educational outcomes or comparing schools. Yet within Greece, not to allude to European or world schools, the results should be interpreted with great caution, since this is not a measuring outcomes study but a tool development study. However, the reported scores could serve as a good starting point for future measurements [65].

5. Conclusion

After a rather successful 'iCAN!v0' debugging, the hereby released 'iCAN!v1' has been created, an acceptably consistent, structured, sensitive (and probably responsive) tool for evaluating and monitoring of whether medical graduates meet the standard set of outcomes produced by the MEDINE network. It discriminates strengths and weaknesses, thus offers a basis for a SWOT educational policy. Any future comprehensive application of it will help obtaining a good picture of medical graduates' competences to inform policies and interventions. Since nobody can prove validity in any absolute sense [46, 63] we expect that any future application will confirm its validity, reliability, sensitivity and responsiveness.

Acknowledgements

The Deans of Greek Medical Schools (at the time of study deployment): C Stefanadis (Athens), G Bourikas (Thrace), N Dombros (Thessaloniki), I Goudevenos (Ioannina) and O Zoras (Crete), for approving this study (the two other Deans, K Malizos and C Gogos, are co-authoring the paper). The schools' secretariat: A Manika and A Vlassi (Athens), A Kirkoudi and V Agrafioti (Thrace), T Gazopoulou and D Nerantzi (Thessaloniki), E Kemalaki (Thessaly), P Sventzouri-Zoi and A Syrou (Ioannina) and E Xenikaki (Crete), for their invaluable help in data collection. Last but not least, the anonymous participants who returned their completed questionnaire.

REFERENCES

- [1] Cumming AD, Ross MT. The Tuning Project for Medicine—learning outcomes for undergraduate medical education in Europe. *Med Teach*. 2007;29(7):636–641.
- [2] Cumming AD, Ross MT. 2008. The Tuning Project (medicine): Learning outcomes / competences for undergraduate medical education in Europe. Edinburgh: The University of Edinburgh [cited 2012 January 14]; Available from: <http://www.tuning-medicine.com>.
- [3] Davis M. 2005. Aims, Objectives and Competences. University of Dundee, Centre for Medical Education, Postgraduate Certificate/ Diploma/ Masters in Medical Education, Theme: Curriculum Development CD4.
- [4] Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C,

- Swing SR et al. Competency-based medical education: theory to practice. *Med Teach.* 2010;32:638-645.
- [5] Guilbert J-J. Educational Handbook for Health Personnel. WHO offset publications No 35. World Health Organization, Geneva 1998. pp 1.36-1.80.
- [6] Harden RM., Crosby JR, Davis MH, Friedman M. AMEE Guide No. 14: Outcome-based education: Part 5. From competency to meta-competency: a model for the specification of learning outcomes. *Med Teach.* 1999;21(6):546-552.
- [7] Harden RM. Developments on outcome based education. *Med Teach.* 2002a;24(2):117-120.
- [8] Harden R.M. Learning outcomes and instructional objectives: is there a difference? *Med Teach.* 2002b;24(2):151-155.
- [9] McGaghie WC, Miller GE, Sajid AW, Telder TV. Competency-based curriculum development in medical education. Geneva: World Health Organization. 1978 [cited 2011 May 6]; Available from: http://whqlibdoc.who.int/php/WHO_PHP_68.pdf.
- [10] Rodgers AT, Tilson ER, Cross DS. Competency, objectives and outcomes. *Radiol Technol.* 2000;71(5):487-490.
- [11] Prideaux D. The emperor's new clothes: from objectives to outcomes. *Med Educ.* 2000;34(3):168-169.
- [12] AMEE 1999. AMEE Medical Education Guide No. 14: Outcome-based Education [cited 2012 January 1]; Available from: <http://www.amee.org/index.asp?lm=102>.
- [13] Harden R.M. Outcome-Based Education: the future is today. *Med Teach.* 2007a;29(7):625-629.
- [14] Davis MH, Amin Z, Grande JP, O'Neill AE, Pawlina W, Viggiano TR, Zuberi R. Case studies in outcome-based education. *Med Teach.* 2007;29(7):717-722.
- [15] Elizondo-Montemayor L, Cid-Garcva A, Pirez-Rodriguez BA, Alarcon-Fuentes G, Pirez-Garcva I, David S. Outcome-based national profile of Mexico's medical graduates. *Med Teach.* 2007;29(7):691-698.
- [16] Ellaway R, Evans P, McKillop J, Cameron HS, Morrison J, McKenzie H et al. Cross-referencing the Scottish Doctor and Tomorrow's Doctors outcome frameworks. *Med Teach.* 2007;29:636-641.
- [17] Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach.* 2007;29(7):642-647.
- [18] Hoat DrLN, Yen NB, Wright EP. Participatory identification of learning objectives in eight medical schools in Vietnam. *Med Teach.* 2007;29(7):683-690.
- [19] Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach.* 2007;29(7):648-654.
- [20] Schwarz MR, Wojtczak A, Stern D. The outcomes of global minimum essential requirements (GME) pilot implementation in China. *Med Teach.* 2007;29(7):699-705.
- [21] Xiao H, Xian L, Yu X, Wang J. Medical curriculum reform in Sun Yat-sen University: implications from the results of GME evaluation in China. *Med Teach.* 2007;29(7):706-710.
- [22] Quinn DC, Bingham JW, Shourbaji NA, Jarquin-Valdivia AA. Medical students learn to assess care using the healthcare matrix. *Med Teach.* 2007;29(7):660-665.
- [23] Core Committee, Institute for International Medical Education. Global minimum essential requirements in medical education. *Med Teach.* 2002;24(2):130-5.
- [24] Stern DT, Wojtczak A, Schwarz MR; IIME Task Force for Assessment. The assessment of Global Minimum Essential Requirements in medical education. *Med Teach.* 2003;25(6):589-95.
- [25] European Communities (2009). ECTS USERS' GUIDE. Education and Culture DG, Brussels, 6 February. Luxembourg: Office for Official Publications of the European Communities. ISBN: 978-92-79-09728-7 doi:10.2766/88064.
- [26] IFMSA/EMSA (International Federation of Medical Students' Associations / European Medical Students' Association). European Core Curriculum- the Students' Perspective. Bristol, UK, 2006 [cited 2010 August 28]; Available from: <http://www.sedem.org/eccsp.pdf>.
- [27] HelMSIC 2006. European Core Curriculum- the Students' Perspective: Greek edition. Bristol, UK [cited 2012 January 16]; Available from: <http://www.scribd.com/doc/24925436/>.
- [28] Berk RA. Thirteen strategies to measure college teaching. Stylus publishing LLC, Sterling, Virginia, USA 2006.
- [29] Lazaridis T. The way of renaissance of the Greek universities. Editions Kritiki, Athens 2008. pp 260 (in Greek): pp 175-178.
- [30] Goldacre MJ, Lambert T, Evans J, Turner G. Preregistration house officers' views on whether their experience at medical school prepared them well for their jobs: national questionnaire survey. *BMJ.* 2003;326:1011-1012 [cited 2010 October 1]; Available from: <http://www.bmj.com/cgi/content/full/326/7397/1011>.
- [31] Eley DS. Postgraduates' Perceptions of Preparedness for Work as a Doctor and Making Future Career Decisions: Support for Rural, Non-traditional Medical Schools. *Educ Health.* 2010;374 [cited 2010 October 10]; Available from: <http://www.educationforhealth.net>.
- [32] Holt MC, Roff S. Development and validation of the Anaesthetic Theatre Educational Environment Measure (ATEEM). *Med Teach.* 2004;26(6):553-558.
- [33] Roff S, McAleer S, Harden RM, Al-Qahtani M, Uddin AA, Deza H et al. Development and Validation of the Dundee Ready Education Environment Measure (DREEM). *Med Teach.* 1997;19(4):295-299.
- [34] Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees (STEEM). *Med Teach.* 2004;26(3):260-264.
- [35] Nagraj S, Wall D, Jones E. The development and validation of the mini-surgical theatre educational environment measure (miniSTEEM). *Med Teach.* 2007;29(6):e192-e197.
- [36] Roff S, McAleer S, Skinner A. Development and Validation of an instrument to measure postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK (PHEEM). *Med Teach.* 2005;27(4):326-331.
- [37] Kanashiro J, McAleer S, Roff S. Assessing the educational

environment in the operating room—a measure of resident perception at one Canadian institution (OREEM). *Surgery*. 2006;139(2):150-158.

- [38] Dimoliatis ID, Vasilaki E, Anastassopoulos P, Ioannidis JP, Roff S. Validation of the Greek translation of the Dundee Ready Education Environment Measure (DREEM). *Educ Health (Abingdon)*. 2010;23:348 [cited 2010 October 16]; Available from: http://www.educationforhealth.net/publish/articles/article_print_348.pdf.
- [39] Rammos A, Tatsi K, Bellos S, Dimoliatis IDK. Translation into Greek of the postgraduate hospital educational environment measure (PHEEM). *Arch Hellen Med*. 2011;28(1):48-56. In Greek with a 297 words English abstract. [cited 2012 January 14]; Available from: <http://www.mednet.gr/archives/2011-1/pdf/48.pdf>.
- [40] Tseretopoulou X, Tzamalís T, Bazoukis G, Dimoliatis I. iCAN! version 1.0: a graduate self-completion questionnaire evaluating medical curriculum outcomes. AMEE 2009 Abstract Book, page S317, abstract 10X8 [cited 2010 October 1]; Available from: <http://www.amee.org/index.asp?pg=94>.
- [41] Tseretopoulou X., Dimoliatis I. 2009. Learning Outcomes and Competences for Undergraduate Medical Education in Europe: The Tuning (medicine) Project. pp26. The Greek translation of Cumming & Ross (2008) [cited 2012 January 15]; Available from: <http://www.tuning-medicine.com/>.
- [42] Whittle S, Whelan B, Murdoch-Eaton DG. DREEM and beyond; studies of the educational environment as a means for its enhancement. *Educ Health (Abingdon)*. 2007;20(1):7 [cited 2010 October 16]; Available from: <http://www.educationforhealth.net/articles/subviewnew.asp?ArticleID=7>.
- [43] Tseretopoulou X, Tzamalís T, Bazoukis G, Lyrakos G, Gogos C, Thermos K, et al. If you could change one thing in your school, what would this be? 469 suggestions of 429 medical graduates. *Arch Hellen Med*. 2011;28 (3):369-379. In Greek with a 333 English abstract [cited 2012 January 14]; Available from: <http://www.mednet.gr/archives/2011-3/369abs.html>.
- [44] Streiner DL, Norman GR. *Health Measurement Scales - a practical guide to their development and use*. 4th edition. Oxford University Press, Oxford 2008.
- [45] Norusis/SPSS Inc. 1997. *SPSS Professional Statistics™ 7.5*. Chapter 13 Measuring scales: reliability analysis examples. pp 103-111.
- [46] Fayers PM, Machin D. *Quality of life: assessment, analysis and interpretation*. John Willey & Sons, England 2006. See also Wikipedia. Anonymous. Spearman-Brown prediction formula http://en.wikipedia.org/wiki/Spearman-Brown_formula.
- [47] Schönrock-Adema J., Heijne-Penninga M., Van Hell E.A. & Cohen-Schotanus J. 2009. Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Med Teach*. 2009;31(6):e226-e232.
- [48] RAND/ Health (2009). Scoring Instructions for MOS 36-Item Short Form Survey Instrument (SF-36). [cited 2012 January 31]; Available from: http://www.rand.org/health/surveys_tools/mos/mos_core_36item.html.
- [49] Pappa E., Kontodimopoulos N. & Niakas D. Psychometric evaluation and normative data for the Greek SF-36 health survey using a large urban population sample. *Arch Hellen Med*. 2006;23(2):159-166. [cited 2008 May 2]; Available from: <http://www.mednet.gr/archives/contents2006-2-en.htm> 1.
- [50] European Commission (2009) SHE figures 2009. Statistics and Indicators on Gender Equality in Science. Directorate General for Research, EUR 23856, European Communities, Luxembourg
- [51] Ruest-Archambault E (2008) Benchmarking Policy Measures for Gender Equality in Science. EUR 23314, European Communities, Belgium.
- [52] Kaldoudi E (2011). Women in Science and Engineering – A Greek Experience. IFMBE News, no. 88, pp. 17-22, Nov-Jan 2011/2012 [cited 2012 August 26]; Available from: <http://www.ifmbe.org/attachments/article/104/IFMBE-NO88-6.zip>.
- [53] Trochim MK. 2006. Regression to the Mean. Web center for social research methods Knowledge base [cited 2012 February 18]; Available from: <http://www.socialresearchmethods.net/kb/regrmean.php>, Last Revised: 20 Oct 2006.
- [54] Reason J. Human error: models and management. *BMJ*. 2000;320:768–770.
- [55] Laufenberg D. (filmed Nov 2010, posted Dec 2010). How to learn? From mistakes. TED ideas worth spreading talks [cited 2012 February 11]; Available from: http://www.ted.com/talks/diana_laufenberg_3_ways_to_teach.html.
- [56] Mamede S, van Gog T, van den Berge K, Rikers R.M.J.P., van Saase J.L.C.M., van Guldener C, Schmidt H.G. Effect of Availability Bias and Reflective Reasoning on Diagnostic Accuracy Among Internal Medicine Residents. *JAMA*. 2010;304(11):1198-1203 [cited 2012 February 18]; Available from: <http://jama.ama-assn.org/content/304/11/1198.full> PubMed ID 20841533.
- [57] Dror I. A novel approach to minimize error in the medical domain: cognitive neuroscientific insights into training. *Med Teach*. 2011;33(1):34-8. PMID 21067318.
- [58] Harford T. (filmed Jul 2011, posted Jul 2011). Trial, error and the God complex. TED ideas worth spreading talks [cited 2012 February 11]; Available from: http://www.ted.com/talks/tim_harford.html.
- [59] Schulz K. (filmed Mar 2011, posted Apr 2011): On being wrong. TED ideas worth spreading talks [cited 2012 February 11]; Available from: http://www.ted.com/talks/kathryn_schulz_on_being_wrong.html.
- [60] Goldman B. (filmed Nov 2011, posted Jan 2012): Doctors make mistakes. Can we talk about that? TED ideas worth spreading talks/TEDx [cited 2012 February 11]; Available from: http://www.ted.com/talks/brian_goldman_doctors_make_mistakes_can_we_talk_about_that.html.
- [61] Tzamalís T, Tseretopoulou X, Bazoukis G, Jelastopulu E, Papadopoulos IN, Dimoliatis IDK. Do medical graduates perceive that their knowledge and abilities result from their school's curriculum? The Greek experience. AMEE 2011 Abstract Book, page 468-469, abstract 10AA13 [cited 2012 January 14]; Available from: <http://www.amee.org/index.asp?pg=206&cookies=True>.
- [62] Dixit H, Marahatta SB. Medical education and training in

- Nepal: SWOT analysis. Kathmandu Univ Med J (KUMJ). 2008;6(23):412-20.
- [63] Norman G, Eva KW. 2008. Quantitative research methods in medical education. Association for the Study of Medical Education booklet series "Understanding Medical Education".
- [64] Tavakol M, Dennick R. 2011. Making sense of Cronbach's alpha. IJME, 2: 53-55.
- [65] Dimoliatis IDK, Tzamalís T, Bazoukis G, Tseretopoulou X, Lyrakos G, Gogos C et. al. iCAN! A graduate self-completion questionnaire for evaluating medical curriculum outcomes: How to use it, and preliminary findings on Greek medical education outcomes. Arch Hellen Med. 2011;28(5):647-666. In Greek with a 456 words English abstract [cited 2012 January 14]; Available from: <http://www.mednet.gr/archives/2011-5/pdf/647.pdf>.
- [66] Norman G. 2010. Likert scales, levels of measurement and the "laws" of statistics. Adv in Health Sci Educ 2010; 15: 625-632.
- [67] Alison Harper on behalf of the WHOQOL group. WHOQOL-BREF: introduction, administration, scoring and generic version of the assessment. Field Trial Version December 1996: page 10.